DOCKET NO.: 262905US0PCT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Pascale GAILLARD et al.

SERIAL NO.: NEW U.S. PCT APPLICATION

FILED: HEREWITH

INTERNATIONAL APPLICATION NO.: PCT/EP03/50225

INTERNATIONAL FILING DATE: June 13, 2003

FOR: AZOLE METHYLIDENE CYANIDE DERIVATIVES AND THEIR USE AS PROTEIN KINASE MODULATORS

REQUEST FOR PRIORITY UNDER 35 U.S.C. 119 AND THE INTERNATIONAL CONVENTION

Commissioner for Patents
Alexandria, Virginia 22313

Sir:

In the matter of the above-identified application for patent, notice is hereby given that the applicant claims as priority:

COUNTRY	APPLICATION NO	DAY/MONTH/YEAR
EPC	02100710.9	14 June 2002

Certified copies of the corresponding Convention application(s) were submitted to the International Bureau in PCT Application No. PCT/EP03/50225.

Respectfully submitted,
OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

Norman F. Oblon
Attorney of Record
Registration No. 24,618
Surinder Sachar
Registration No. 34,423

Customer Number
22850

(703) 413-3000
Fax No. (703) 413-2220
(OSMMN 08/03)

PCT/EP 03 / 50225

Europäisches Patentamt
European Patent Office
Office européen des brevets

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

02100710.9

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Europäisches Patentamt
European Patent Office
Office européen des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 02100710.9
Demande n°:

Anmeldetag:
Date of filing: 14/06/02
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
Applied Research Systems ARS Holding N.V.
Curacao
NETHERLANDS ANTILLES

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
Azole methylidene cyanide derivatives and their use as protein kinase modulators

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat: State: Pays:	Tag: Date: Date:	Aktenzeichen: File no. Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:
C07D417/06, C07D417/14, A61K31/506, A61P3/10

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE/TR
Etats contractants désignés lors du dépôt:

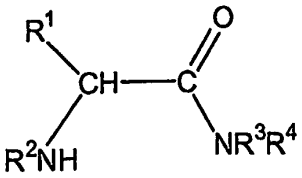
Bemerkungen:
Remarks:
Remarques:
See for the original title of the application, page 1 of the description

POLYPEPTIDES HAVING α-H-α-AMINO ACID AMIDE RACEMASE ACTIVITY AND NUCLEIC ACIDS ENCODING THE SAME.

The invention relates to isolated polypeptides having α-H-α-amino acid amide racemase activity and nucleic acids encoding the same. The invention also relates to vectors and host cells comprising the nucleic acids according to the invention. The invention also relates to methods of producing and using the polypeptides according to the invention. The invention also relates to a method for isolating polypeptides having α-H-α-amino acid amide racemase activity, for isolating nucleic acids encoding the same and for isolating microorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity. The invention also relates to new microorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity.

α-H-α-Amino acid amides are readily available compounds and are important precursors in the production of pharmaceuticals and of α-H-α-amino acids. For example, enantiomerically enriched α-H-α-amino acids can be obtained from mixtures of D- and L-α-H-α-amino acid amides with randomly chosen enantiomeric excess (ee) by enantioselective enzymatic hydrolysis of one of the enantiomers of the α-H-α-amino acid amide. In such a process for the preparation of enantiomerically enriched α-H-α-amino acids, simultaneous racemization of the α-H-α-amino acid amides would be of great advantage because then complete conversion of the α-H-α-amino acid amide into the desired optically active α-H-α-amino acid is possible. Also, in other processes racemization of α-H-α-amino acid amides is often desired. As enzymatic racemization is preferable over chemical racemization (mild reaction conditions, environmental benefits etc.), many attempts have been made to identify microorganisms with α-H-α-amino acid amide racemase activity and to isolate polypeptides with α-H-α-amino acid amide racemase activity and genes encoding this activity.

For the purpose of the present invention, α-H-α-amino acid amide racemase activity is defined as the ability to catalyze the racemization of an enantiomerically enriched α-H-α-amino acid amide according to formula 1,



(1)

wherein R^1 stands for an optionally substituted alkyl of 1-20 C-atoms or an optionally substituted (hetero)aryl of 3-20 C-atoms (carbon atoms of the substituents included) and wherein R^2, R^3 and R^4 each independently stand for H or an optionally substituted alkyl of 1-20 C-atoms or an optionally substituted (hetero)aryl of 3-20 C-atoms (carbon atoms of the substituents included) and wherein R^1 may form a ring with R^2 and/or R^3 may form a ring with R^4 together with the carbon and/or nitrogen atoms to which they are bound. Each ring is preferably 5-8 membered. Preferably, R^1, R^2, R^3 or R^4 each independently stand for an alkyl of 1-10 C-atoms or a (hetero)aryl of 3-10 C-atoms (carbon atoms of the substituents included). Substituens on the alkyl or (hetero) aryl may be chosen from the group of: hydroxy, alkoxy, mercapto, thioalkyl, alkyl, carboxy, amino, imino, nitro, halogen, carbamoyl, cyano, acyl, peroxy, sulpho or phospho. Examples of α-H-α-amino acid amides according to formula1 are: proline amide, leucine amide, glutamic acid amide, phenylalanine amide, t-leucine amide, methionine amide, tryptophan amide, leucyl-glycine, valyl-glycine, valyl-alanine, leucyl-alanine.

With α-H-α-amino acid amide racemase is meant a polypeptide having α-H-α-amino acid amide racemase activity.

α-H-α-Amino acid amide racemase activity can be detected with methods similar to methods employed to determine racemizing activity of racemases acting on compounds other than α-H-α-amino acid amides. Such methods are known to the person skilled in the art. For example, by measuring the decrease in enantiomeric excess of an L-α-H-α-amino acid amide or of a D-α-H-α-amino acid amide, α-H-α-amino acid amide racemase activity can be determined.

In Fukumura *et al*, 1977, Agric. Biol. Chem., vol. 41, p 1509-1510, an α-amino-e-caprolactam racemase isolated from *Achromobacter obae* is disclosed; the nucleic acid sequence coding for this enzyme and the amino acid sequence of this enzyme were published in Naoko *et al.*, 1987, Biochemistry of vitamin B6, p 449-452. For its enzymatic activity, the α-amino-e-caprolactam racemase isolated from *Achromobacter obae* requires pyridoxal-5-phosphate as a cofactor (Ahmed *et al.*, 1983, Agric. Biol. Chem., vol 47, p 1887-1893).

The α-amino-ε-caprolactam racemase isolated from *Achromobacter obae* is known to exclusively catalyze the racemization of α-amino-ε-caprolactam. Although the α-amino-ε-caprolactam racemase from *Achromobacter obae* was tested for α-H-α-amino acid amide racemase activity with the following substrates, which are α-H-α-amino acid amides according to formula 1: L-tryptophan amide, L-leucine amide, L-leucyl-glycine, L-valyl-glycine, L-valyl-L-alanine, L-leucyl-L-alanine, no α-H-α-amino acid amide racemase activity was found even when an excess amount of enzyme was used (Ahmed *et al.*, 1983, Agric. Biol. Chem., vol 47, p 1887-1893).

EP-A-383403 discloses α-H-α-amino acid amide racemase activity in the genus *Klebsiella* and related genera and EP-B1-378592 discloses α-H-α-amino acid amide racemase activity in *Arthrobacter sp.* ATCC 31652 (DSM 4639) and in *Corynebacterium sp.* ATCC 31662 (DSM 4640). However, using the microorganisms disclosed in these publications, no racemization could be detected. Therefore, the documents EP-A-383403 and EP-B1-378592 should not be considered as prior art.

So, although much work has been done to find an α-H-α-amino acid amide racemase, up till now, not one α-H-α-amino acid amide racemase has been disclosed.

The invention now provides such α-H-α-amino acid amide racemases.

Applicants surprisingly also discovered a suitable method to isolate micoorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity, in which a microorganism containing sample is enriched by using D-α-amino-ε-caprolactam or a mixture of D- and L-α-amino-ε-caprolactam as the sole nitrogen source and in which the thus found cultures are tested for α-H-α-amino acid amide racemase activity.

A suitable way for carrying out such a method is for example as follows; this method will be further referred to as the enrichment method. To find polypeptides according to the invention a microorganism containing sample, e.g. an environmental sample, for example a soil sample or a waste water sample is cultured in or on a suitable growth medium containing D-α-amino-ε-caprolactam or a mixture of D- and L-α-amino-ε-caprolactam as the sole nitrogen source until growth can be detected.

The sample can be directly added onto/into the suitable growth medium with D-α-amino-ε-caprolactam or a mixture of D- and L-α-amino-ε-caprolactam as the sole nitrogen source. There are also other ways of applying the sample to the medium, for example by adding the filtrate of a wash solution used to wash the sample

with. Growth temperatures are preferably adapted to the temperatures of the natural environment of the microorganisms present in the samples.

The culturing of the environmental sample in or on a suitable growth medium containing D-α-amino-e-caprolactam or a mixture of D- and L-α-amino-e-caprolactam as the sole nitrogen source is a so-called enrichment and was described by Fukumura *et al.*, 1977, Agric. Biol. Chem., vol 41, p 1321-1325, who used this enrichment in a method to isolate α-amino-e-caprolactam racemases. Preferably, the enrichment is continued by one or more transfers of the cultured microorganism(s) into or onto a 'fresh' suitable growth medium containing D-α-amino-e-caprolactam or a mixture of D- and L-α-amino-e-caprolactam as the sole nitrogen source until a monoculture is reached. Typically, this will be after 4 or 5 transfers.

The microorganisms obtained by enrichment are tested for α-H-α-amino acid amide racemase activity. This testing for α-H-α-amino acid amide racemase activity can be done directly on the whole cells or on permeabilized cells of the colonies obtained after plating the cultured microorganisms. Alternatively, the colonies obtained are cultured separately in a suitable growth medium containing D-α-amino-e-caprolactam or a mixture of D- and L-α-amino-e-caprolactam as the sole nitrogen source, after which cell free extract is prepared therefrom, which is subsequently tested for α-H-α-amino acid amide racemase activity.

Cell free extract can be prepared according to standard methods known to the person skilled in the art, e.g. by sonification, French press etc. Cell permeabilization can be obtained according to standard methods known to the person skilled in the art, e.g. by addition of small amounts of toluene. Suitable growth media are in fact all media, which do not contain a nitrogen source next to D-α-amino-e-caprolactam or a mixture of D- and L-α-amino-e-caprolactam. Suitable growth media are well known in the art. They can for example be composed by the person skilled in the art with guidance from Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual.2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Surprisingly, it was found that the enrichment method is suitable for the isolation of microorganisms displaying α-H-α-amino acid amide racemase activity. This is the more surprising since this enrichment selects for α-amino-e-caprolactam racemase activity and not for α-H-α-amino acid amide racemase activity.

Several microorganisms might be isolated this way, for example it may be possible to isolate microorganisms from the following genera: *Agrobacterium,*

Ochrobactrum, Arthrobacter, Micrococcus, Aureobacterium, Corynebacterium, Rhodococcus, Brevibacterium, Rubrobacter, Nocardioides, Terrabacter.

Several monocultures of strains of microorganisms displaying α-H-α-amino acid amide racemase activity were identified with the enrichment method and these microorganisms were deposited under the Budapest Treaty with The National Collections of Industrial and Marine Bateria Limited (NCIMB), Aberdeen, Scotland, on May 8th, 2002: *Agrobacterium rhizogenes* Na was deposited under number NCIMB 41127, *Agrobacterium rhizogenes* Bi was deposited under number NCIMB 41128, *Arthrobacter nicotianae* was deposited under number NCIMB 41126, *Ochrobactrum anthropi* 1A was deposited under number NCIMB 41129.

Therefore, the invention also relates to a method for isolating microorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity, comprising the steps of:

a) Culturing, in one or more transfer steps, a microorganisms containing sample in or on a suitable growth medium containing D-α-amino-ε-caprolactam or a mixture of D- and L-α-amino-ε-caprolactam as the sole nitrogen source
b) Testing the thus obtained microorganisms for α-H-α-amino acid amide racemase activity.

Testing for α-H-α-amino acid amide racemase activity is preferably performed on a monoculture, but can also be performed on a mix culture. The invention also relates to new microorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity obtainable with said method.

In a preferred embodiment of the invention, the mix and/or monocultures obtained after a), are tested for the ability to use both L- and D-α-amino-ε-caprolactam as the sole nitrogen source.

From microorganisms obtained with the enrichment method it is possible to isolate a nucleic acid sequence encoding a polypeptide having α-H-α-amino acid amide racemase activity and to clone and express this nucleic acid to produce a polypeptide according to the invention. The isolation of a nucleic acid sequence and subsequent recloning and expression thereof can be done according to standard methods, which are known to the person skilled in the art. An example of such a nucleic acid isolation method is as follows.

In a first step total DNA is isolated and a gene library is prepared from the microorganisms obtained by the enrichment method and expressed in a suitable

vector in a suitable host (e.g. as described in the examples). In a second step, the clones of the gene library containing a vector with insert (the insert is a piece of DNA isolated from the microorganisms) are tested on the ability to catalyse the racemization of an α-H-α-amino acid amide. For example, testing of the clones for α-H-α-amino acid amide racemase activity can be done according to the methods as described in the examples.

In subsequent steps, the nucleic acid sequence of the insert of the vector in a clone having the ability to catalyse the racemization of an α-H-α-amino acid amide is determined and open reading frames are identified from the thus determined nucleic acid sequence, after which the open reading frames are recloned and expressed in a suitable vector and in a suitable host and again tested for α-H-α-amino acid amide racemase activity. By expression of the open reading frame displaying α-H-α-amino acid amide racemase activity in a suitable vector in a suitable host a polypeptide having α-H-α-amino acid amide racemase activity according to the invention can be produced.

With the enrichment method and nucleic acid isolation method the nucleic acid sequence encoding an α-H-α-amino acid amide racemase from a strain identified as *Ochrobactrum anthropi* 1A deposited under number NCIMB 41129 with the NCIMB was obtained. This nucleic acid sequence is presented in SEQ ID: NO. 1. The amino acid sequence of the corresponding polypeptide is presented in SEQ ID: NO. 2.

Therefore, the invention also relates to a method for isolating a nucleic acid sequence encoding a polypeptide with α-H-α-amino acid amide racemase activity, comprising the steps of the enrichment method (steps a and b) and which method further comprises the steps of:

c) Isolating the total DNA from the microorganism displaying α-H-α-amino acid amide racemase activity and cloning the isolated total DNA into a suitable vector into a suitable host to produce and express a gene library
d) Testing the clones from the gene library for α-H-α amino acid amide racemase activity
e) Isolating the vectors with insert from the clones displaying α-H-α-amino acid amide racemase activity
f) Determining the nucleic acid sequence of the inserts in the vectors of the clones displaying α-H-α-amino acid amide racemase activity
g) Identifying open reading frames from the sequenced inserts.
h) Recloning and expressing the open reading frames in suitable vectors in suitable hosts and testing the hosts for α-H-α-amino acid amide racemase

activity.

The invention also relates to nucleic acid sequences encoding polypeptides with α-H α-amino acid amide racemase activity obtainable by the above method. The invention also relates to a method for producing a polypeptide with α-H-α-amino acid amide racemase activity, comprising the steps a-h and which method further comprises the step of:

i) Expressing the open reading frame in a suitable vector in a suitable host to produce a polypeptide with α-H-α-amino acid amide racemase activity.

The invention also relates to polypeptides with α-H-α-amino acid amide racemase activity obtainable by the above method.

In a preferred embodiment of the invention, the invention relates to isolated polypeptides having α-H-α-amino acid amide racemase activity and having a degree of identity with the amino acid sequence presented in SEQ ID: NO. 2 of at least about 55%, preferably at least about 65%, more preferably at least 75%, even more preferably at least 85%, in particular at least 90%, more in particular at least 95% and most in particular at least 97%.

For purpose of the present invention, the degree of identity between two amino acid sequences is determined by the blastp pairwise alignment algorithm (NCBI) with an identity table and the following alignment parameters: mismatch = -3 penalty = -3, gap extend = 1, match bonus = 1, Gap x – droff = 50, expect = 10, wordsize = 3.

The present invention also relates to isolated polypeptides having α-H-α-amino acid amide racemase activity, which are encoded by nucleic acid sequences which hybridize under low stringency conditions, preferably under medium stringency conditions, more preferably under high stringency conditions and most preferably under very high stringency conditions with the coding sequence of SEQ ID: NO. 1 or a complementary strand thereof.

Hybridization experiments can be performed by a variety of methods, which are well available to the skilled man. General guidelines for choosing among these various methods can be found in e.g. chapter 9 of Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual.2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

With stringency of the hybridization conditions is meant, the conditions under which the hybridization, consisting of the actual hybridization and wash steps, are performed. Wash steps are used to wash off the nucleic acids, which do not hybridize

with the target nucleic acid immobilized on for example a nitrocellulose filter. The stringency of the hybridization conditions can for example be changed by changing the salt concentration of the wash solution and/or by changing the temperature under which the wash step is performed (wash temperature). Stringency of the hybridization increases by lowering the salt concentration in the wash solution or by raising the wash temperature. For purpose of this application, the hybridization is performed in 6 X sodium chloride/sodium citrate (SSC) at about 45°C for about 12 hours. Two consecutive 30 minutes wash steps in 1 X SSC, 0.1% SDS at 50°C is an example of low stringency, at 55°C an example of medium stringency, at 60°C an example of high stringency, at 65°C an example of very high stringency.

The present invention also relates to isolated polypeptides having α-H-α-amino acid amide racemase activity and which display immunological cross-reactivity with an antibody raised against a fragment of the amino acid sequence according to SEQ ID: NO. 2.

The immunological cross reactivity may be assayed using an antibody raised against, or reactive with, at least one epitope of the isolated polypeptide according to the present invention having α-H-α-amino acid amide racemase activity. The antibody, which may either be monoclonal or polyclonal, may be produced by methods known in the art, e.g. as described by Hudson *et al.*, Practical Immunology, Third Edition (1989), Blackwell Scientific Publications. The immunochemical cross-reactivity may be determined using assays known in the art, an example of which is Western blotting, e.g. as described in Hudson *et al.*, Practical Immunology, Third Edition (1989), Blackwell Scientific Publications.

The invention also relates to fragments of the polypeptides according to the invention and having α-H-α-amino acid amide racemase activity of at least 100 amino acids, preferably of 125 to 350 amino acids, more preferably of 200 to 300 amino acids.

The invention also relates to fusion proteins made by expression of a nucleic acid sequence encoding a polypeptide according to the invention operatively linked to one or more nucleic acid sequences, which encode (a) marker polypeptide(s). With operatively linked is meant, that the nucleic acid sequences are linked such that, if expressed, the polypeptide according to the invention with the marker polypeptide(s) on its N- and/or C-terminus is produced. The marker polypeptide can serve many purposes, for example, it may be used to increase the stability or the solubility of the fusion protein, it may be used as a secretion signal, which is a signal that directs the fusion protein to a

certain compartment in the cell or it may be used to facilitate purification of the fusion protein. An example of a marker polypeptide used to facilitate purification of the fusion protein is the hexahistidine peptide. The purification of a fusion protein with a hexahistidine tag is for example described in Gentz *et al.*, 1989, Proc. Natl. Acad. Sci. USA, vol. 86, p. 821-824. A fusion protein with a hexahistidine tag can for example be produced in a pQE vector (Qiagen, Inc.), by following the protocol of the supplier.

The invention also relates to nucleic acid sequences encoding polypeptides having α-H-α-amino acid amide racemase activity according to the invention. These nucleic acid sequences can be obtained with the enrichment method and nucleic acid isolation method as described above.

A nucleic acid sequence of the present invention, such as a nucleic acid sequence with the sequence of SEQ ID: NO. 1 can also be isolated using standard molecular biology techniques and the sequence information provided herein. For example, using all or a portion of the nucleic acid sequence of SEQ ID: NO. 1 as a hybridization probe, a nucleic acid sequence according to the invention can be isolated using standard hybridization and cloning techniques (e. g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual.2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid sequence encompassing all or a portion of SEQ ID: NO. 1 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence information contained in SEQ ID: NO. 1 or SEQ ID: NO. 2.

A nucleic acid sequence of the invention can be amplified using for example genomic DNA, cDNA or alternatively mRNA as a template and appropriate oligonucleotide primers according to standard (RT)-PCR amplification techniques. The nucleic acid so amplified can be cloned into a suitable vector and characterized by DNA sequence analysis.

Furthermore, oligonucleotides corresponding to or hybridizable to nucleic acid sequences according to the invention can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

After isolation of a nucleic acid sequence encoding a polypeptide according to the invention, the nucleic acid can be cloned in a suitable vector and after introduction in a suitable host, the sequence can be (over)expressed according to standard cloning and expression techniques, which are known to the person skilled in the art (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular

Cloning: A Laboratory Manual.2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), to produce a polypeptide according to the invention. Therefore, the invention also relates to vectors comprising a nucleic acid sequence according to the invention.

Alternatively, to produce a polypeptide according to the invention, the nucleic acid sequence encoding a polypeptide according to the invention can be integrated into the genome of a host cell and be (over)expressed. This can be done according to methods known to the person skilled in the art. Therefore, the invention also relates to a host cell comprising a nucleic acid sequence according to the invention, preferably to a host cell comprising a vector comprising a nucleic acid sequence according to the invention.

Alternatively, a polypeptide according to the invention can be overexpressed in its natural host, for example by placing a suitable promoter upstream of the nucleic acid sequence according to the invention, by integrating one or more copies of a nucleic acid sequence according to the invention into the genome of its natural host, preferably by expressing a nucleic acid according to the invention in a suitable vector in its natural host, which are all methods known to the person skilled in the art.

Suitable vectors are the vectors normally used for cloning and expression and are known to the person skilled in the art. Examples of suitable vectors for expression in *E. coli* are given e.g. in table 1 in Makrides, S.C., 1996, Microbiological Reviews, p. 512-538. Preferably, the vector contains a promoter upstream of the cloning site containing the nucleic acid sequence encoding the polypeptide with α-H-α-amino acid amide racemase activity, which can be switched on after the host has been grown to express the corresponding polypeptide having α-H-α-amino acid amide racemase activity. Promoters, which can be switched on and off are known to the person skilled in the art and are for example the *lac* promoter, the *ara*BAD promoter, the T7 promoter, the *trc* promoter, the *tac* promoter, the *trp* promoter, and the *aro*H promoter. Suitable hosts are the hosts normally used for cloning and expression and are known to the person skilled in the art. Examples of suitable *E. coli* host strains are: TOP10F', TOP10, DH10B, DH5α, HB101, W3110, BL21(DE3) and BL21(DE3)pLysS.

The choice of the vector can sometimes depend on the choice of the host and vice versa. If e.g. a vector with the *ara*BAD promoter is being used, an *E. coli* host strain that is unable to break down the arabinose inducer (*ara*⁻ host strain) is strongly preferred.

It should be noted that the invention also includes mutants of the nucleic acids encoding polypeptides having α-H-α-amino acid amide racemase activity, which have one or more mutations as compared to the corresponding polypeptide isolated from a naturally occurring host. Methods for making mutations are known to the person skilled in the art, e.g. by random mutagenesis (for example by error prone PCR or by UV radiation), site directed mutagenesis, etc.

α-H-α-Amino acid amide racemases according to the invention can suitably be used together with an enantioselective amidase in a process for the preparation of enantiomerically enriched α-H-α-amino acids from a mixture of D- and L-α-H-α-amino acids amides or in a process for the preparation of L-α-H-α-amino acids from the corresponding D-α-H-α-amino acid amides or in a process for the preparation of D-α-H-α-amino acids from the corresponding L-α-H-α-amino acid amides. Use of α-H-α-amino acid amide racemases in such processes will lead to a 100% conversion of the α-H-α-amino acid amide into the corresponding enantiomerically enriched α-H-α-amino acid (100% yield).

Preferably, in the above-mentioned processes, the enantioselective amidase is at least 90% enantioselective, more preferably at least 95% enantioselective, even more preferably at least 98% enantioselective, most preferably at least 99% enantioselective. 90% Enantioselectivity of the amidase (for example a D-amidase) is defined for the present invention as the ability of the amidase to convert a racemic mixture of DL-α-H-α-amino acid amides into 90% of one enantiomer of the α-amino acid (e.g. D-α-H-α-amino acid) and into 10% of the other enantiomer of the α-H-α-amino acid (e.g. L-α-H-α-amino acid) at 50% total conversion of the α-H-α-amino acid amide mixture. 95% Enantioselectivity (e.g. 95% D- and 5% L-α-H-α-amino acid) corresponds to an E-value of 58.4 (e.e. 90%). 98% Enantioselectivity (e.g. 98% D- and 2% L-α-H-α-amino acid) corresponds to an E-value of 193.6 (e.e. 96%). 99% Enantioselectivity (e.g. 99% D- and 1% L-α-H-α-amino acid) corresponds to an E-value of 458.2 (e.e. 98%).

The invention therefore also relates to the use of an α-H-α-amino acid amide racemase according to the invention in combination with an enantioselective amidase in a process for the preparation of enantiomerically enriched α-H-α-amino acids from a mixture of D- and L-α-H-α-amino acids amides or in a process for the preparation of L-α-H-α-amino acids from the corresponding D-α-H-α-amino acid amides or in a process for the preparation of D-α-H-α-amino acids from the corresponding L-α-H-α-amino acid amides.

Preferably, if used in said processes, the α-H-α-amino acid amide racemase according to the invention does not display α-H-α-amino acid racemase activity, i.e. the ability to catalyse the racemization of a D-α-H-α-amino acid and an L-α-H-α-amino acid.

The invention is illustrated by way of the following examples. However, these examples are not meant to restrict the invention.

Examples

Example 1

Enrichment method

In this enrichment procedure, different soil and sludge samples that had been stored at -80°C, were used for direct inoculation of 50 ml of liquid Medium A (Yeast Carbon Base (Difco, 11.7 g/l), KH_2PO_4 (6.7 g/l), K_2HPO_4 (8.9 g/l), pH 7.0), containing racemic DL-α-amino-ε-caprolactam (2 g/l) as sole nitrogen source. Per flask, one spatula of inoculum was used. The different flasks were incubated at 28°C and 200 rpm.

After about 2 days good growth was observed in all flasks. After sedimentation of the soil particles in all cultures, 1:1,500 dilutions were prepared in Medium A containing 2 g/l DL-α-amino-ε-caprolactam, and were incubated at 28°C and 200 rpm. After obtaining sufficient growth, samples were taken from each flask to determine the concentration of both L- and D-α-amino-ε-caprolactam in the culture broths. These samples were centrifuged and the supernatants were filtered through a 0.45 µM filter to remove all cells. Then these samples were analyzed by HPLC. o-Phthaldehyde in combination with D-3-mercapto-2-methylpropionic acid was used in this HPLC method as a chiral reagent for the separation of both α-amino-ε-caprolactam enantiomers (Duchateau *et al.*, 1992, J. Chromatogr., vol. 623, p 237 – 245).

Only flasks in which cultures had grown that used both enantiomers of α-amino-ε-caprolactam were selected for the next steps. Different dilutions of these cultures were plated onto plates with Medium A containing 2 g/l DL-α-amino-ε-caprolactam as sole nitrogen source, and were incubated at 28°C. Colonies were randomly chosen and reisolated on the same type of selective plates to obtain monocultures. Then, a colony of each monoculture was transferred into 5 ml of Medium A containing 2 g/l DL-α-amino-ε-caprolactam. After 3 days of incubation at 28°C and 200 rpm, the cultures obtained were analyzed for use of D- and L-α-amino-ε-caprolactam

exactly as described above. The monocultures that could use both D- and L-α-amino-e-caprolactam were stored at -80°C in 20% (v/v) glycerol.

Cells from these monocultures were finally tested for α-amino-e-caprolactam racemase activity. After cultivation in Medium A containing DL-α-amino-e-caprolactam (2 g/l) at 28°C and 200 rpm for three days, cells were harvested by centrifugation, washed with 20 mM HEPES-NaOH buffer (pH 7.7) and lyophilized.

Then 100 mg of these lyophilized cells from each monoculture were added to a reaction mixture consisting of 10 mM HEPES-NaOH buffer (pH 7.7), toluene (2,5 v/v %), pyridoxal-5-phosphate (20 µM), and L-α-amino-e-caprolactam (2.5 m%). Chemical blanks (with the assay mixture without the lyophilized cells) were used as a negative control. All reaction mixtures (including the blanks) were incubated for approximately 72 hours at 40°C before the reaction was stopped by the addition of 9 volumes of 1 M H_3PO_4. After removal of the cells by filtration through a 0.45 µM filter, these reaction mixtures were analyzed by the HPLC method described above to determine the concentration of D- and L-α-amino-e-caprolactam. Finally, monocultures that formed D-α-amino-e-caprolactam from the L-α-amino-e-caprolactam substrate were sent to the NCIMB for strain determination.
Nb. no D-α-amino-e-caprolactam could be determined in the chemical blanks.

Strain determination

One of the thus obtained monocultures that could racemize L-α-amino-e-caprolactam was identified as an *Ochrobactrum anthropi* strain by the NCIMB and was deposited under the Budapest Treaty as *Ochrobactrum anthropi* IA with the NCIMB under number NCIMB 41129. The results of the determination are presented in table 1.

Table 1: Results of strain identification of isolate IA by National Collections of Industrial and Marine Bacteria Limited (NCIMB LTD). Identification number ID4036.

Isolate code	IA
Incubation temperature (°C)	30
Gram stain	-
Spores	-
Motility	+
Colony morphology (after 2 days growth on LB medium)	Round, regular, entire, smooth, shiny, low convex, buff, opaque. 1 mm diameter.
Growth at 37°C	+
Growth at 45°C	-
Catalase	+
Oxidase	+
Fermentative in glucose OF	Oxidative
Results from API 20NE test:	
NO_3 reduction	+
Indole reduction	-
Acid from glucose	-
Arginine dihydroxylase	-
Urease	+
Aesculin hydrolysis	-
Gelatin hydrolysis	-
β-galactosidase	-
Glucose assimilation	+
Arabinose assimilation	+
Mannose assimilation	+
Mannitol assimilation	+
N-acetyl-glucosamine assimilation	+
Maltose assimilation	-
Gluconate assimilation	-
Caprate assimilation	-
Adipate assimilation	-
Malate assimilation	+
Citrate assimilation	+

Isolate code	IA
Phenylacetate assimilation	-
Cytochrome oxidase	+
Results from further biochemical tests:	
Acid production in:	
Glucose ASS[a]	+
Dulcitol ASS	+
Adonitol ASS	+
Raffinose ASS	-
Xylose OF	+
Glycine CSU[b]	+
Nitrite reduction to N_2	+
PPA	+
Simmons citrate	-
Tween 80 hydrolysis	-
Tyrosine decomposition	-
H_2S production (PbAc)	+

[a]: ASS, ammonia salt sugar.

[b]: CSU, carbon source utilization.

Measurement of α-H-α-amino acid amide racemase activity of lyophilized *Ochrobactrum anthropi* IA cells

O. anthropi IA cells were cultivated in Medium A containing DL-α-amino-e-caprolactam (2 g/l) as sole nitrogen source. After incubation at 28°C and 200 rpm for three days ($OD_{620\ nm}$ = 4.7), cells were harvested by centrifugation, washed in 20 mM HEPES-NaOH buffer of pH 7.7 and lyophilized.

The lyophilized cells were used to demonstrate α-H-α-amino acid amide racemase activity. The assay mixture consisted of 10 mM HEPES-NaOH buffer (pH 7.7) containing toluene (2,5 v/v %), pyridoxal-5-phosphate (20 µM), D- or L-leucine amide (2.5 m%) and in some cases EDTA (20 mM). Per assay 150 mg of lyophilized cells were used. Chemical blanks (with the assay mixture without the *O. anthropi* IA cells) were used as a negative control. All reaction mixtures (including the blanks) were incubated for approximately 72 hours at 40°C before the reaction was stopped by the addition of 9 volumes of 1 M H_3PO_4. Finally, the reaction mixtures were analyzed by HPLC to determine the concentrations of D- and L-leucine amide and of D- and L-

leucine. *o*-Phthaldehyde in combination with D-3-mercapto-2-methylpropionic acid was used in this HPLC method as a chiral reagent for the enantioseparation of these amino compounds (Duchateau *et al.*, 1992, J. Chromatogr., vol. 623, p 237 – 245).

In a number of reactions EDTA was added to inhibit L-and/or D-specific amidase(s) present in the *O. anthropi* IA cells. By inhibiting this amidase, the L- and/or D-amino acid amide is not or just partly converted into the corresponding amino acid thereby enabling the demonstration of amino acid amide racemase activity by the detection of the other amino acid amide enantiomer. These reactions with EDTA (No. 3 & 4), clearly showed that *O. anthropi* IA cells contain an α-H-α-amino acid amide racemase activity, because D-leucine amide was converted into L-leucine amide and vice versa. The chemical blanks did not show this racemization reaction.

Without EDTA, amidases convert the substrate and/or product α-H-α-amino acid amides into α-H-α-amino acids thereby preventing direct detection of the other α-H-α-amino acid amide enantiomer as racemization product. Starting from D-leucine amide, however, the reaction without EDTA also clearly proved the presence of an α-H-α-amino acid amide racemase activity in *O. anthropi* IA, because this substrate was converted to significant amounts of L-leucine, and a control experiment excluded its formation via D-leucine, because *O. anthropi* IA did not racemize leucine. Formation of L-leucine from D-leucine amide could not be detected in the chemical blank.

Table 2: α-H-α-amino acid amide racemase activity in lyophilized *O. anthropi* IA cells.

No.	Substrate	EDTA (20 mM)	D-leucine amide (m%)	L-leucine amide (m%)	D-leucine (m%)	L-leucine (m%)
1	D-leucine amide	-	0.22	<0.001	0.015	0.044
2	L-leucine amide	-	<0.001	<0.001	0.001	0.21
3	D-leucine amide	+	0.24	0.032	0.007	0.008
4	L-leucine amide	+	0.021	0.15	0.003	0.058

The concentrations given are the concentrations measured in the reaction mixtures, which are 10 times diluted as compared to the initial reaction mixtures.

Example 2

Enrichment method

Approximately 30 g of fresh soil samples from different locations were resuspended in 50 ml of Medium A, supplemented with 2 g/l D-α-amino-e-caprolactam as sole nitrogen source, and shaken for 30 minutes at 4°C and 200 rpm. The suspensions were filtered through a filter paper to remove the soil particles. With the filtrates obtained the following two different approaches were used to select for strains that could use D-α-amino-e-caprolactam as nitrogen source:

- 5 ml Of the different filtrates were transferred to empty 100 ml flasks. All flasks were incubated at 28°C and 200 rpm until an $OD_{600\ nm}$ was reached of approximately 2.5. Then the cultures were diluted 1:100 in fresh Medium A containing 2 g/l D-α-amino-e-caprolactam, again followed by a cultivation step. These dilution and recultivation step was repeated 3 more times. Then different dilutions of good growing cultures were plated onto plates with Medium A containing 2 g/l D-α-amino-e-caprolactam as sole nitrogen source, that were incubated at 28°C. Colonies were randomly chosen and reisolated 5 times onto identical selective plates, yielding pure monocultures. Finally, these monocultures were stored at –80°C in 20% (v/v) glycerol.
- As alternative enrichment strategy, different dilutions of all filtrates were plated directly onto Medium A containing 2 g/l D-α-amino-e-caprolactam as sole nitrogen source. After 2 days of incubation at 28°C, first colonies appeared on almost all plates. After a few more days to 3 weeks, a lot of morphologically different colonies were obtained. Morphologically different colonies were randomly chosen and reisolated 5 times onto identical selective plates, yielding pure monocultures. Finally, these monocultures were stored at –80°C in 20% (v/v) glycerol.

Cells from these monocultures were finally tested for α-amino-e-caprolactam racemase activity. After cultivation in 800 ml of Medium A containing 2 g/l of D-α-amino-e-caprolactam at 28°C and 200 rpm, cells were harvested by centrifugation (20 min. at 5,000 x g, 4°C), and resuspended in 40 ml of sonification buffer (NaCl, 16 g/l; KCl, 0.74 g/l; Na_2HPO_4, 0.27 g/l; glucose, 2 g/l; HEPES, 10 g/l, pH 7.0). Then the cells were disintegrated by sonification using a Soniprep 150 of MSE with a 0.6 mm nozzle at an amplitude of 20 microns in cycles of 10 min. using time intervals of 10 sec. (i.e. 10 sec. sonification, 10 sec. cooling) and cooling in ice/acetone. After each cycle, the cells were viewed microscopically and the sonification procedure was stopped when 50 - 70% of the cells were broken.

Then D-α-amino-ε-caprolactam and pyridoxal-5-phosphate were added to the disintegrated cell suspensions of each monoculture to concentrations of 5 g/l and 0.01 mM respectively. After 1.5 h of incubation at 20°C and 20 rpm, 2 ml samples were transferred into 1 ml of 1 M H_6PO_4. After removal of all particulates by centrifugation followed by filtration through a 0.22 µM filter, concentrations D- and L-α-amino-ε-caprolactam in the samples were determined by HPLC using the protocol as described in example 1. Finally, monocultures that formed L-α-amino-e-caprolactam from the D-α-amino-e-caprolactam substrate were sent to the NCIMB for strain determination.

Strain determination

The thus obtained monocultures were identified by the NCIMB as *Agrobacterium rhizogenes* (*Agrobacterium rhizogenes* Bi and *Agrobacterium rhizogenes* Na and *Arthrobacter nicotianae* via 16S rDNA sequence determination. The results of this 16S rDNA determination are presented below. The microorganisms were deposited under the Budapest Treaty with the NCIMB. *Agrobacterium rhizogenes* Na was deposited under number NCIMB 41127, *Agrobacterium rhizogenes* Bi under number NCIMB 41128 and *Arthrobacter nicotianae* under number NCIMB 41126.

Analysis of 16S rDNA sequence

Methods

DNA extraction: The DNA was extracted using the Prepman purification kit and stored on ice until use.

Polymerase chain reaction: The 16S ribosomal DNA gene was amplified using universal eubacterial primers and analysed by electrophoresis on a 1% agarose gel.

DNA clean-up: The 500 bp fragment was purified by spin column centrifugation and resuspended in sterile destilled water.

DNA sequencing: The purified DNA product was automatically sequenced using the dideoxy chain terminator method (Sanger *et al.*, 1997).

Sequence analysis by database comparison: The sequence of the 16S rDNA gene was compared with nucleic acid sequence databases.

Results

The 16S rDNA sequences of *Agrobacterium rhizogenes* Na NCIMB 41127 (SEQ ID: NO. 5), *Agrobacterium rhizogenes* Bi NCIMB 41128 (SEQ ID: NO. 4),

and *Arthrobacter nicotianae* NCIMB 41126 (SEQ ID: NO. 3) are presented in the sequence listing part.

Measurement of α-H-α-amino acid amide racemase activity of disintegrated cells of *Agrobacterium rhizogenes* Na, *Agrobacterium rhizogenes* Bi, *Arthrobacter nicotianae*

Preparation of disintegrated cells

After growth of these three strains in Medium A containing 2 g/l of D-α-amino-ε-caprolactam, cells were harvested by centrifugation at 12.000 x g. After centrifugation the supernatant was decanted and the wet weight of the pellet was measured. To the pellet, sonification buffer (NaCl, 16 g/l; KCl, 0.74 g/l; Na_2HPO_4, 0.27 g/l; glucose, 2 g/l; HEPES, 10 g/l, pH 7.0) was added in a 1:2 ratio (wet weight cell pellet : ml buffer). During storage between cell harvest and sonification the pellet was frozen in the sonification buffer at –86°C.

Sonification was performed with 0.6 mm nozzle in a Soniprep 150 of MSE at maximal amplitude of 20 microns in periods of 10 min. using time intervals of 10 sec. (i.e. 10 sec. sonification, 10 sec. cooling down), cooling the sample on ice/acetone. Each 10 min. the cells were viewed microscopically. The sonification procedure was stopped when >70% of the cells were broken.

α-H-α-Amino acid amide racemase activity test with disintegrated cells

1 ml Of disintegrated cells of each of the three strains was incubated in a HEPES-NaOH buffer (20 mM) pH 7.7, to which pyridoxal-5-phosphate (0.01 mM) was added as a cofactor and D-leucine amide (60.8 mM) as a substrate in a total volume of 10 ml (0.79 w/w% D-leucine amide). After 0 and 139 hours 1 ml samples were taken (of which the exact weight was measured) and the reaction in these samples was stopped by adding 1 ml methanol (of which the exact weight was measured too; to be able to calculate dilution factors). Stopped reaction mixtures were centrifuged to remove particulates. The supernatants were frozen at -86°C until HPLC analysis according to the method given in example 1.

Blank reaction mixtures were prepared, sampled and stopped in the same way, with the exception that no substrate was added and that instead of disintegrated cells, only HEPES-NaOH buffer (20 mM, pH 7.7) was added.

The results are shown in table 3.

Table 3. Measurement of α-H-α-amino acid amide racemase activity of disintegrated cells of *Agrobacterium rhizogenes* Na, *Agrobacterium rhizogenes* Bi and *Arthrobacter nicotianae*.

Sample code	Incubation time (h)	D-leucine (w/w%)	L-leucine (w/w%)	D-leucine amide (w/w%)
A. nicotianae	139	0.002	0.068	0.56
A. rhizogenes Na	139	0.036	0.078	0.51
A. rhizogenes Bi	139	0.013	0.025	0.61

From the table it can be seen that α-H-α-amino acid amide racemase activity is present in the disintegrated cells of *Agrobacterium rhizogenes* Na, *Agrobacterium rhizogenes* Bi and *Arthrobacter nicotianae* as D-leucine amide is converted into L-leucine; this conversion is not detected in the chemical blanks. In the chemical blanks no D-leucine or L-leucine could be detected even after 139 hours; the D-leucine amide concentration was constant (0.7 w/w%) meaning that D-leucine amide is very stable under the applied reaction conditions.

Example 3

Isolation of the α-H-α-amino acid amide racemase gene from *Ochrobactrum anthropi* IA.

Expression library construction

To obtain single colonies, a glycerol stock of *O. anthropi* IA was streaked onto a yeast carbon base plate containing 0.5% (w/v) DL-α-amino-ε-caprolactam as single nitrogen source, and cultivated at 28°C. A single colony was transferred to 150 ml of LB liquid medium, and grown at 28°C with vigorous shaking to an $OD_{620\ nm}$ of 0.9. Then cells were harvested by centrifugation and frozen at -20°C.

After thawing the cells, total DNA was isolated according to the Qiagen Genomic DNA Purification Procedure for bacteria using Qiagen Genomic-tip 100/G tips. The standard Qiagen protocol was followed with the following modifications:

- The first incubation step to effect cell lysis was performed without proteinase K for 2 h at 37°C, whereafter twice the suggested amount of proteinase K was added and the solution was incubated for another 2 h at 50°C.
- Before application of the lysate to the Qiagen tips, a centrifugation step (10 min. at 5,000 x g, 4°C) was applied to pellet the particulate matter.

• The lysate from the cells of the 150 ml culture was applied to 3 G/100 columns. 50 µg of the obtained chromosomal DNA was then partially digested with *Sau*3A I at 1/12 U per µg DNA for 30 minutes. Half of the digested DNA was run on a 0.6% agarose gel and DNA fragments between 4 and 10 kb in size were isolated and redissolved in 20 µl of 10 mM Tris-HCl, pH 8.0 buffer.

Vector DNA was prepared by the digestion of 1 µg of pZErO-2 (Invitrogen, Groningen, The Netherlands) with *Bam*H I according to the protocol of Invitrogen.

Linearized vector DNA (50 ng) and *O. anthropi* IA chromosomal DNA fragments (10 µl) were ligated with T4 DNA ligase (according to the protocol of Invitrogen). Subsequently, the ligation mixture was precipitated with NaAc/ethanol and resuspended in 20 µl TE-buffer. 1 µl of this solution was used for the electroporation of electrocompetent *E. coli* DH10B cells (Life Technologies) using a BioRad gene pulser (conditions: 2.5 kV, 25 µF, 100 Ω, 0.1 cm cuvette, 40 µl cell suspension) according to the protocol of the supplier. Transformants were plated onto LB medium with 50 mg/l kanamycin and incubated at 28°C for 24 h. In total over 12,000 colonies were obtained which formed the primary gene library. All 12,000 colonies were pooled in LB medium supplemented with 50 mg/l kanamycin. Part of the cell suspension obtained was used for a total plasmid isolation according to the QiaPrep procedure (Qiagen). This "library in plasmid form" was stored at –20°C till further use. To the remaining part of the cell suspension, glycerol was added to a final concentration of 20% (v/v). The resulting suspension was stored in aliquots at -80°C as primary gene library.

Preparation of L-aminopeptidase help solution

The L-aminopeptidase help solution for the α-H-α-amino acid amide racemase screening procedure was prepared from a recombinant *E. coli* strain containing plasmid pTrcLAP. The *E. coli* expression vector pTrcLAP contains the *Pseudomonas putida* ATCC 12633 *pep*A gene under the control of the *trc* promoter. Detailed information on this *P. putida* L-aminopeptidase encoding gene can be found in T. Sonke, B. Kaptein, W. H. J. Boesten, Q. B. Broxterman, H. E. Schoemaker, J. Kamphuis, F. Formaggio, C. Toniolo, F. J. T.Rutjes in *Stereoselective Biocatalysis* (Ed.: R. N. Patel). Dekker, New York. 2000, pp. 23-58.

A fresh overnight culture of *E. coli* TOP10 / pTrcLAP was used to inoculate 200 ml of LB medium containing both 0.4 mM of IPTG and 100 mg/l of ampicillin. After overnight growth at 30°C, cells were pelleted by centrifugation and

resuspended in 4 ml of 50 mM Tris-HCl, pH 7.5. After cell disruption by one passage through a french press (pressure of 140 Mpa in a 4 ml french press cell), solid particles were collected by centrifugation (45 min. at 40,000 x g, 4°C). The pellet was resuspended in 2 ml of Tris-HCl, pH 7.5 containing 100 mM of $MgSO_4$. This suspension was gently stirred for 30 min. at 4°C. After removal of the particles via centrifugation (45 min. at 40,000 x g, 4°C), the clear solution was stored in aliquots at –20°C for use in the screening assay.

Screening

The "library in plasmid form" solution was diluted 1,000 times in water. 1 µl of this plasmid solution was used to transform electrocompetent *E. coli* TOP10 cells (Invitrogen), and the transformants were plated onto LB medium with 50 mg/l kanamycin. After 2 days of incubation at 28-30°C the obtained colonies were large enough to be transferred to 200 µl of liquid screening medium (tryptone 16 g/l; yeast extract 3 g/l; NaCl 5 g/l; glycerol 2 g/l; pH 7.3) with 50 mg/l kanamycin in microtiter plates. The cultures were grown for 2 days at 28-30°C. Next, a replica of the microtiter plates was prepared by the transfer of a few microliter of each well to new microtiter plates containing solid (0.8% agar) LB medium with 50 mg/l kanamycin. These plates were incubated at 28-30°C for 16-20 h, after which they were stored at 4°C as masterplates. The cells from the remaining part of the cell suspensions were harvested by centrifugation (10 min. at 1,500 x g, room temperature), washed twice with 50 mM Tris-HCl buffer, pH 7.5, and finally re-suspended in 50 µl of 50 mM Tris-HCl buffer, pH 7.5.

The screening reaction was started by the addition of 50 µl of substrate solution containing a mixture of 140 mM D-phenylalanine amide, D-leucine amide and D-valine amide each in 50 mM Tris-HCl, pH 7.5 and 2 mM $MnCl_2$. After 20 h incubation of the microtiterplates on a shaker at 30°C, 2 µl of L-aminopeptidase help solution (for preparation see earlier section) was added. After an additional 1.5 h incubation at 30°C, all reactions were stopped by the addition of 100 µl of 0.15 M HCl.

Subsequently, the ammonium concentration in all reaction mixtures was determined by transfer of 7 µl of these mixtures to new microtiter plates containing 93 µl of GDH reagent per well. This GDH reagent contained per 100 ml 43 mg of NADH, 116 mg of α-ketoglutaric acid, 11.8 mg of ADP and 1200 U of glutamate dehydrogenase (Sigma) in 150 mM Tris-HCl, pH 8.0. After 15 min. incubation at 37°C the $OD_{340\ nm}$ in each well was measured using a Spectramax plus microtiter plate reader (Molecular Devices, Sunnyvale, California, USA). Clones that showed an $OD_{340\ nm}$ that was lower than the mean value of the microtiter plate containing this clone decreased with three

times the standard deviation of that same microtiter plate, were regarded as potential positive clones.

Of 11,272 clones screened, 32 clones could be identified as potential positive clones.

Confirmation of potential positive clones

Of all 32 potential positive clones identified in the screening, material from the masterplates was used to inoculate 3 ml of liquid screening medium containing 50 mg/l kanamycin. After growth for 2 days at 30°C, cells were collected by centrifuging 2 ml of these cell cultures. Then the cell pellets were washed four times in 25 mM Tris-HCl buffer, pH 7.5. Subsequently, the cell pellets were resuspended in 100 µl substrate solution containing 70 mM of D-phenylalanine amide, or D-leucine amide, or D-valine amide in 50 mM Tris-HCl, pH 7.5 and 1 mM $MnCl_2$. After incubation for 20 h at 30°C, the reactions were stopped by the addition of 100 µl of 0.15 M HCl, after which the reaction mixtures were analyzed by the HPLC method described in example 1.

Of the 32 potential positive clones identified in the screening, one showed significant formation of L-leucine from D-leucine amide. This clone contained the *O. anthropi* IA α-H-α-amino acid amide racemase gene on its 7.7 kb plasmid. This was concluded from the fact that transformation of this plasmid into *E. coli* TOP10 cells inevitably led to recombinant cells that converted D-leucine amide into L-Leucine. Sequencing of this plasmid, that was named pOa(1)PLV49B10, revealed the complete nucleotide sequence of the α-H-α-amino acid amide racemase gene. The sequence of this gene is listed as nucleotides 98 to 1417 (including the TAA stopcodon) of SEQ ID: NO. 1 encoding the protein of SEQ ID: NO. 2 as presented *infra*.

Example 4: Construction of plasmid pKEC-AZAR

The *O. anthropi* IA α-H-α-amino acid amide racemase gene was subcloned into *E. coli* expression vector pKECaroP using PCR. Expression vector pKECaroP is similar to construct pKECtrp, whose construction has been described in WO 00/66751, except that pKECaroP contains the pSC101 derived par function and the *E. coli aro*H promoter instead of the *trp* promoter. The α-H-α-amino acid amino racemase open reading frame was amplified using

5'- GCCTCACATATGCAAACACCGCTTTCATTGCG - 3' [SEQ ID: NO. 6]

as forward primer (with *Nde* I cleavage site underlined), and

5'- GCCTCACCCGGGTTACCACATCATAAAATGGGCGACATC - 3' [SEQ ID: NO. 7]

as reverse primer (with *Xma* I cleavage site underlined), and plasmid pOa(1)PLV49B10 as template. This PCR, that was performed with PCR SuperMix High Fidelity (Life Technologies) according to the supplier's protocol, yielded a single fragment. Correct size (1,341 bp) of the amplified fragment was confirmed by agarose gel electrophoresis.

After purification of the amplified fragment with the QIAquick PCR Purification Kit (Qiagen), the fragment was cloned into vector pCR®4Blunt-TOPO (Invitrogen). The cloning mix was subsequently used to transform One Shot™ Chemically Competent *E. coli* TOP10 Cells (Invitrogen). Recombinant cells were selected by plating the whole transformation mixture on LB plates containing 100 µg/ml carbenicillin, followed by overnight incubation at 28°C.

After overnight cultivation of material from six colonies in 5 ml LB medium containing 100 µg/ml carbenicillin, plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (Qiagen). Digestion with restriction enzymes *Nde* I and *Xma* I proved that five out of these six colonies contained the desired recombinant vector.

Plasmid DNA of the five correct clones was pooled and digested to completion with *Nde* I and *Xma* I. The total digestion mixture was applied to a preparative 1% agarose gel for separation of the different fragments. The correct fragment (1,322 bp) was subsequently isolated from the gel by the QIAquick Gel Extraction Kit (Qiagen) and stored at –20°C till further use as insert fragment.

Plasmid pKECaroP was digested to completion with *Nde* I and *Xma* I, yielding two fragments of 2,850 and 3,036 bp as shown by analytical agarose gel electrophoresis. After heat inactivation of the *Nde* I and *Xma* I restriction enzymes (20 minutes, 65°C), *Bsa* I was added to the mixture to cut the undesired 2,850 bp fragment into two smaller pieces. The complete digestion mixture was loaded on a 1% preparative agarose gel, followed by isolation of the desired 3,036 bp fragment using the QIAquick Gel Extraction Kit (Qiagen), that was stored at –20°C till further use as vector fragment.

Vector and insert fragment were ligated using T4 DNA Ligase and the resulting ligation mixture was used for transformation of One Shot™ Chemically Competent *E. coli* TOP10 Cells. The transformation mix was plated on LB plates containing 50 µg/ml kanamycin, that were incubated at 28°C till sufficiently large colonies appeared.

Colony-PCR using PCR SuperMix (Life Technologies) and the above given primers ([SEQ ID: NO. 6] and [SEQ ID: NO. 7]) was performed to screen for colonies containing the correct recombinant vector. Material from twelve PCR positives was used to inoculate 12 tubes with 5 ml LB medium containing 50 µg/ml kanamycin. Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit. Digestion with restriction

enzyme *Hind* III yielded two fragments of 1,731 and 2,627 bp with all twelve plasmids, proving that all twelve colonies contained the desired recombinant vector.

Five of these twelve positive clones were cultured in 25 ml LB medium supplemented with 50 µg/ml kanamycin. After overnight incubation at 28°C, cells were harvested via centrifugation and washed in the same sonification buffer as used in example 2. After resuspending the cell pellets in sonification buffer (ratio wet weight cells to sonification buffer 1:10), the cells were disintegrated by sonification. Crude cell extracts were obtained by removal of the particulates via centrifugation.

The five obtained crude cell extracts were tested for α-H-α-amino acid amide racemase activity by mixing 0.5 ml thereof with 4.5 ml substrate solution consisting of a 22.2 mM HEPES-NaOH buffer pH 7.7 containing 0.01 mM of pyridoxal-5-phosphate and 66 mM D-leucine amide. Reactions were incubated at room temperature. Samples were taken after 0, 18 and 44 hours that were transferred to an equal volume of 1 M H_3PO_4 to stop the reaction. These samples were analyzed by HPLC according to the method given in example 1.

With the crude extract of two of the five clones, called *E. coli*/pKEC_AZAR_3 and *E. coli*/pKEC_AZAR_11, significantly more L-leucine and L-leucine amide was detected in the 18 and 44 hour samples than in the 0 hour sample. Finally, nucleotide sequencing of pKEC_AZAR_3 and pKEC_AZAR_11 revealed the correct nucleotide sequence of the α-H-α-amino acid amide racemase gene and flanking vector parts in these two recombinant plasmids.

Example 5

α-H-α-Amino acid amide racemase activity test using D-leucine amide as substrate and whole *E. coli* cells containing the α-H-α-amino acid amide racemase gene from *Ochrobactrum anthropi* IA as biocatalyst

Preparation of the cells

A single colony of *E. coli* DH10B containing plasmid pOA(1)PLV49B10 was transferred to LB medium with kanamycin (50 mg/l) to prepare a pre-culture. After overnight incubation at 28°C and 200 rpm, this pre-culture was used to inoculated a flask with 500 ml of 2*TY medium (10 g/l Yeast extract, 16 g/l Tryptone, 5 g/l NaCl) containing kanamycin (50 mg/l). After overnight growth at 28°C (OD_{620nm} = 4.4) and 200 rpm, cells were harvested by centrifugation (15 min. at 6,200 x g, 4°C), and washed with 50 mM HEPES-NaOH buffer of pH 7.7.

To be able to determine the background activity of the *E. coli* host-vector system in the activity tests, an *E. coli* DH10B strain containing a pZErO-2 based construct with a mutant Green Fluorescent Protein (GFPuv) encoding gene as insert in the opposite direction as the vector borne *lac* promoter, was cultivated via the same procedure (*E. coli* DH10B / pZErO-GFPuv-wrong-orientation). Growth of this recombinant *E. coli* strain resulted in an overnight culture with an OD_{620nm} of 3.8.

The cell pellets from both cultures were subsequently washed and resuspended in 25 ml of 50 mM HEPES-NaOH buffer, pH 7.7. Aliquots of 2 ml of both cell suspensions were centrifuged once more and the pellets were stored at -20°C until execution of the activity tests described below. The remaining part of these cell suspensions was stored at –20°C for use in example 6.

<u>α-H-α-amino acid amide racemase activity towards D-leucine amide</u>

To determine the activity of *E. coli* / pOA(1)PLV49B10 and the *E. coli* control, the above-mentioned cell pellets were thawed and resuspended to a total volume of 1 ml with 50 mM HEPES-NaOH buffer, pH 7.7. Then, reaction mixtures of 10 ml each were prepared containing 50 mM HEPES-NaOH buffer (pH 7.7), 10 µM pyridoxal-5-phosphate (PLP), 2.5 wt% of D-leucine amide, if applicable 20 mM of EDTA to suppress the *E. coli* amidase activity, and 0.5 ml of the *E. coli* / pOA(1)PLV49B10 and *E. coli* DH10B / pZErO-GFPuv-wrong-orientation cell suspensions or water (for chemical blanks). Reactions were started by the addition of the cells. Reaction mixtures were incubated at 30°C and 175 rpm. Directly after addition of the cells (t = 0 hours) and after 27 and 43 hours samples were taken in which the reaction was stopped by removal of the cells via centrifugation followed by 0.22 µM filtration. Finally, the filtered samples were stored at –20°C until analysis by chiral HPLC as described below:

column:	Sumichiral OA5000 from Sumika (150 x 4.6 mm I.D., 5µ) + guard column
eluent:	85 v/v% 2mM $CuSO_4$ + 15 v/v% methanol
flow:	1.0 ml/min.
column temp.:	40°C
inj. volume:	5µl
detection:	fluorescence detection after post-column reaction with *o*-phthalaldehyde and 2-mercaptoethanol (wavelength ex=338 nm and em>420 nm)

The results of this experiment are presented in table 4.

Table 4. α-H-α-amino acid amide racemase activity towards D-Leucine amide of *E. coli* DH10B / pOA(1)PLV49B10 and *E. coli* blank (*E. coli* DH10B / pZErO-GFPuv-wrong-orientation).

Strain	20 mM EDTA	Incubation time (h)	D-leucine amide (wt%)	L-leucine amide (wt%)	D-leucine (wt%)	L-leucine (wt%)
B	-	0	2.67	n.d.	n.d.	n.d.
B	-	27	2.51	n.d.	0.003	0.005
B	-	43	2.63	0.060	0.005	0.004
A	-	0	2.63	n.d.	n.d.	n.d.
A	-	27	1.03	0.170	0.009	1.71
A	-	43	0.512	0.063	0.017	2.12
A	+	0	2.61	n.d.	n.d.	n.d.
A	+	27	1.77	0.555	0.003	0.313
A	+	43	1.59	0.772	0.006	0.345

n.d.: Not detectable

Strain A: *E. coli* DH10B / pOA(1)PLV49B10

Strain B: *E. coli* DH10B / pZErO-GFPuv-wrong-orientation

The data in table 4 clearly show that *E. coli* DH10B / pZErO-GFPuv-wrong-orientation could not convert D-leucine amide. Even after 43 hours of reaction, the reaction mixture contained the same amount of D-leucine amide as at the start of the reaction. This was also the case for all chemical blanks (data not shown).

With *E. coli* DH10B / pOA(1)PLV49B10 cells on the other hand, the amount of D-leucine amide clearly decreased in time. Without additional EDTA, this substrate was converted to a relatively low amount of L-leucine amide and a large amount of L-leucine. With additional EDTA, a much higher concentration of L-leucine amide was obtained, because EDTA, a chelating compound partially inhibits the amidase activity of *E. coli*, thereby reducing the conversion of L-leucine amide to L-leucine.

The results obtained in this experiment prove that *E. coli* DH10B / pOA(1)PLV49B10 cells contain α-H-α-amino acid amide racemase activity towards D-leucine amide. Furthermore, they show that by combining this α-H-α-amino acid

amide racemase with an L-selective amidase / aminopeptidase (as present in e.g. *E. coli* DH10B), D-α-H-α-amino acid amides can be converted to L-α-H-α-amino acids.

Example 6

α-H-α-Amino acid amide racemase activity test using DL-leucine amide as substrate and whole *E. coli* cells containing the α-H-α-amino acid amide racemase gene from *Ochrobactrum anthropi* IA as biocatalyst

The *E. coli* DH10B / pOA(1)PLV49B10 and *E. coli* DH10B / pZErO-GFPuv-wrong-orientation cell suspensions from example 5 (Preparation of the cells) were used. To determine their activity towards racemic DL-leucine amide, identical reaction mixtures were prepared as in example 5, except that 2.5 wt% of DL-leucine amide was used, EDTA was ommitted in all reactions, and the reactions were started with 2 ml of the cell suspensions. Samples were taken directly after the start of the reactions (t = 0 hours) and after 8 and 24 hours.

The results of this experiment are presented in table 5.

Table 5. α-H-α-amino acid amide racemase activity towards DL-leucine amide of *E. coli* DH10B / pOA(1)PLV49B10 and *E. coli* blank (*E. coli* DH10B / pZErO-GFPuv-wrong-orientation).

Sample code	Incubation time (h)	D-leucine amide (wt%)	L-leucine amide (wt%)	D-leucine (wt%)	L-leucine (wt%)	$e.e._{L\text{-leucine}}$ (%)
B	0	1.13	1.15	0.001	0.004	-
B	8	1.15	0.12	0.001	1.43	99.9
B	24	1.23	n.d.	0.002	1.48	99.7
A	0	1.19	1.25	0.001	0.003	-
A	8	1.24	0.33	0.004	1.45	99.4
A	24	0.43	0.14	0.028	2.25	97.5

n.d. : Not detectable

Strain A: *E. coli* DH10B / pOA(1)PLV49B10

Strain B: *E. coli* DH10B / pZErO-GFPuv-wrong-orientation

From the data in table 5 it becomes clear that the *E. coli* blank cells (*E. coli* DH10B / pZErO-GFPuv-wrong-orientation) can only convert L-leucine amide to L-leucine with an enantiomeric excess of over 95%. The D-leucine amide is left

untouched, thereby resulting in a maximum yield of 50% only.

With the *E. coli* DH10B / pOA(1)PLV49B10 cells containing the α-H-α-amino acid amide racemase encoding gene, clearly both L- and D-leucine amide are converted to L-leucine, leading to a yield of over 50% and maximum 100%. Again, the enantiomeric excess of the obtained L-leucine is well over 95%.

This experiment proves that by combining the α-H-α-amino acid amide racemase from *O. anthropi* IA with an L-selective amidase / aminopeptidase (as present in e.g. *E. coli* DH10B), DL-α-H-α-amino acid amides can be converted into L-α-H-α-amino acids in a more than 50% yield.

CLAIMS

1. Isolated polypeptide having α-H-α-amino acid amide racemase activity and having a degree of identity with the amino acid sequence presented in SEQ ID: NO. 2 of at least about 55%, preferably of at least about 65%, more preferably 75%, even more preferably 85%, in particular 90%, more in particular 95%, most in particular 97%.

2. Isolated polypeptides having α-H-α-amino acid amide racemase activity, which are encoded by nucleic acid sequences which hybridize under low stringency conditions, preferably under medium stringency conditions, more preferably under high stringency conditions, most preferably under very high stringency conditions with SEQ ID: NO. 1 or a complementary strand thereof.

3. Isolated polypeptides having α-H-α-amino acid amide racemase activity which display immunological cross reactivity with an antibody raised against a fragment of the amino acid sequence according to SEQ ID: NO. 2.

4. Isolated polypeptides or fragments thereof of at least 100 amino acids, preferably 125-350 amino acids, more preferably 200-300 amino acids, with α-H-α-amino acid amide racemase activity.

5. Isolated fusion protein made by expression of a nucleic acid sequence encoding a polypeptide according to any of claims 1-4 operatively linked to one or more nucleic acid sequences, which encode (a) marker polypeptide(s).

6. Nucleic acid sequence encoding a polypeptide according to any of claims 1-4 or a fusion protein according to claim 5.

7. Vector comprising a nucleic acid sequence according to claim 6.

8. Host cell comprising a nucleic acid sequence according to claim 6 or a vector according to claim 7.

9. Use of a polypeptide according to any of claims 1-4 or a fusion protein according to claim 5 together with an enantioselective amidase in a process for the preparation of enantiomerically enriched α-H-α-amino acids from a mixture of D- and L-α-H-α-amino acid amides or in a process for the preparation of L-α-H-α-amino acids from the corresponding D-α-H-α-amino acid amides or in a process for the preparation of D-α-H-α-amino acids from the corresponding L-α-H-α-amino acid amides.

10. Method for isolating microorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity, comprising the steps of:
    a) Culturing, in one or more transfer steps, a microorganism containing sample in or on a suitable growth medium containing D-α-amino-ε-caprolactam or a mixture of D- and L-α-amino-ε-caprolactam as the sole nitrogen source;
    b) Testing the thus obtained microorganisms for α-H-α-amino acid amide racemase activity.

11. Microorganisms obtainable by the method of claim 10.

12. Method for isolating a nucleic acid sequence encoding a polypeptide with α-H-α-amino acid amide racemase activity, comprising the steps of claim 10 and further comprising the steps of:
    c) Isolating the total DNA from the microorganisms displaying α-H-α-amino acid amide racemase activity and cloning the isolated total DNA into a suitable vector into a suitable host to produce and express a gene library,
    d) Testing the clones from the gene library for α-H-α amino acid amide racemase activity,
    e) Isolating the vectors with insert from the clones displaying α-H-α-amino acid amide racemase activity,

f) Determining the nucleic acid sequences of the inserts in the vectors of the clones displaying α-H-α-amino acid amide racemase activity,

g) Identifying open reading frames from the sequenced inserts,

h) Recloning and expressing the open reading frames in suitable vectors in suitable hosts and testing the hosts for α-H-α-amino acid amide racemase activity.

13. Nucleic acid encoding a polypeptide with α-H-α-amino acid amide racemase activity obtainable by the method of claim 12.

14. Method for the production of a polypeptide with α-H-α-amino acid amide racemase activity, comprising the step of:

i) expressing the identified open reading frame obtained with the method of claim 12 in a suitable vector in a suitable host to produce a polypeptide with α-H-α-amino acid amide racemase activity.

15. Polypeptide with α-H-α-amino acid amide racemase activity obtainable by the method of claim 14.

ABSTRACT

The invention relates to isolated polypeptides having α-H-α-amino acid amide racemase activity and nucleic acids encoding the same. The invention also relates to vectors and host cells comprising the nucleic acids according to the invention. The invention also relates to methods of producing and using the polypeptides according to the invention. The invention also relates to a method for isolating polypeptides having α-H-α-amino acid amide racemase activity, for isolating nucleic acids encoding the same and for isolating microorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity. The invention also relates to new microorganisms comprising polypeptides having α-H-α-amino acid amide racemase activity.

Sequence Listing

<110> DSM N.V.

<120> POLYPEPTIDES HAVING alpha-H-alpha-AMINO ACID AMIDE RACEMASE ACTIVITY AND NUCLEIC ACIDS ENCODING THE SAME

<130> V20595

<160> 7

<170> PatentIn version 3.1

<210> 1

<211> 1501

<212> DNA

<213> Ochrobactrum anthropi IA NCIMB 41129

<220>

<221> CDS

<222> (98)..(1417)

<223> Amino Acid Amide Racemase encoding sequence

<400> 1

attcgtcagg atcttgtatc gacgccacag ggcatagccg aatggacaga tcgtctggca 60

aaggctttgc ttggtcagtc acaccagggg tgatgtg atg caa aca ccg ctt tca 115
Met Gln Thr Pro Leu Ser
1 5

ttg cgt gag cgc gac gca cgg gtc att gca gaa ata gga cgc tta cgg 163
Leu Arg Glu Arg Asp Ala Arg Val Ile Ala Glu Ile Gly Arg Leu Arg
10 15 20

ttt tca cct ttg agc ctg atc ggc ggc aag ggc aac agg ctg ata gag 211
Phe Ser Pro Leu Ser Leu Ile Gly Gly Lys Gly Asn Arg Leu Ile Glu
25 30 35

gaa ggt ggg cgt tcc atc ctt gat ctg tct ggc tcc gcc gga cct gct 259
Glu Gly Gly Arg Ser Ile Leu Asp Leu Ser Gly Ser Ala Gly Pro Ala
40 45 50

gct ctg ggg tat ggt cat ccg gct att gtt gaa gcg gtt gaa aaa tct 307
Ala Leu Gly Tyr Gly His Pro Ala Ile Val Glu Ala Val Glu Lys Ser
55 60 65 70

gtt cgc gat atg gcg ggt gca agc ttg ttg ctt tac ccg aac gag gct 355
Val Arg Asp Met Ala Gly Ala Ser Leu Leu Leu Tyr Pro Asn Glu Ala
75 80 85

gcc gtt tcg ctt gcc gaa gac ttg ctg cgc atc acc ccc ggc aat ggc 403
Ala Val Ser Leu Ala Glu Asp Leu Leu Arg Ile Thr Pro Gly Asn Gly
90 95 100

gag cgt cgc gtg tgg ttt ggt cat tcg gga tca gat gcc aat gat tgt 451
Glu Arg Arg Val Trp Phe Gly His Ser Gly Ser Asp Ala Asn Asp Cys

105 110 115

gcg gta cgc gta cta act gcc gcc acc aag cgc tca cgt atc ata tct 499
Ala Val Arg Val Leu Thr Ala Ala Thr Lys Arg Ser Arg Ile Ile Ser
120 125 130

ttt atc ggt tct tac cac ggt aat ctg aca ggc tcg atg ggg ata agc 547
Phe Ile Gly Ser Tyr His Gly Asn Leu Thr Gly Ser Met Gly Ile Ser
135 140 145 150

ggt cac act gcg atg aca cat aca ttg ccg cgc ccc ggt gta ttg ctg 595
Gly His Thr Ala Met Thr His Thr Leu Pro Arg Pro Gly Val Leu Leu
155 160 165

ttg cca tat cct gat ccg ttc agg ccg cgc ttt tct gcg gaa gcc gtg 643
Leu Pro Tyr Pro Asp Pro Phe Arg Pro Arg Phe Ser Ala Glu Ala Val
170 175 180

ctt gaa ctt ctc gat tat cac ttc gct aca agc tgc ccg cca gag caa 691
Leu Glu Leu Leu Asp Tyr His Phe Ala Thr Ser Cys Pro Pro Glu Gln
185 190 195

gtt gcg gct gtg ttc atc gag ccg atc tta tca gat ggt ggt ctc gtc 739
Val Ala Ala Val Phe Ile Glu Pro Ile Leu Ser Asp Gly Gly Leu Val
200 205 210

gtg cct ccg ccg gcc ttt tta gaa gca ttg cag gat cgt tgt cgc aag 787
Val Pro Pro Pro Ala Phe Leu Glu Ala Leu Gln Asp Arg Cys Arg Lys
215 220 225 230

cat gga att ctg gtg gta gtc gat gag gtc aag gtt ggc ctt ggc cga 835
His Gly Ile Leu Val Val Val Asp Glu Val Lys Val Gly Leu Gly Arg
235 240 245

acc ggg ctg atg cat tgt ttc cag cat gaa ggg ctt gag cct gac atg 883
Thr Gly Leu Met His Cys Phe Gln His Glu Gly Leu Glu Pro Asp Met
250 255 260

gtg gtg ttc gga aaa ggt ctt gga ggt ggt tta cct ctt tct gcg gtt 931

```
Val Val Phe Gly Lys Gly Leu Gly Gly Gly Leu Pro Leu Ser Ala Val
        265                 270                 275

gtt gga ccg caa tgg gtg atg gat cac gcg cct gcg ttt gtg ttg caa      979
Val Gly Pro Gln Trp Val Met Asp His Ala Pro Ala Phe Val Leu Gln
    280                 285                 290

acc aca gcc gga aat cca gtg gcg act gct gcc ggt cgc gct gta ctc     1027
Thr Thr Ala Gly Asn Pro Val Ala Thr Ala Ala Gly Arg Ala Val Leu
295                 300                 305                 310

aat acg att gag agg caa gga ctt gcc caa cga tcg gag cgg gtt ggg     1075
Asn Thr Ile Glu Arg Gln Gly Leu Ala Gln Arg Ser Glu Arg Val Gly
                315                 320                 325

ggg att ttc gca gat cgc ctg cgc cgg ctc agc gat aag cat tcc atc     1123
Gly Ile Phe Ala Asp Arg Leu Arg Arg Leu Ser Asp Lys His Ser Ile
            330                 335                 340

ata ggg gat gta cgc ggc aga ggg cta gcg att ggc gtt gat ctc gta     1171
Ile Gly Asp Val Arg Gly Arg Gly Leu Ala Ile Gly Val Asp Leu Val
        345                 350                 355

agt gat cgt gga tca cgt gag ccg gct ccg gta acg acg acc gcg aaa     1219
Ser Asp Arg Gly Ser Arg Glu Pro Ala Pro Val Thr Thr Thr Ala Lys
    360                 365                 370

att ata tat cgc ggc tat caa ctc ggc gca gct ttt acc tat gtc ggc     1267
Ile Ile Tyr Arg Gly Tyr Gln Leu Gly Ala Ala Phe Thr Tyr Val Gly
375                 380                 385                 390

ctc aat gcc aat gtt ctg gaa ttc atg ccc ccg ttg act ttg acc gag     1315
Leu Asn Ala Asn Val Leu Glu Phe Met Pro Pro Leu Thr Leu Thr Glu
                395                 400                 405

ccg gaa ata gac gaa gca gcg gac atc gtt gat cag gcc att ggc gat     1363
Pro Glu Ile Asp Glu Ala Ala Asp Ile Val Asp Gln Ala Ile Gly Asp
            410                 415                 420
```

gtt ctg gat ggg aag gtg gct gat agc gat gtc gcc cat ttt atg atg 1411
Val Leu Asp Gly Lys Val Ala Asp Ser Asp Val Ala His Phe Met Met
425 430 435

tgg taa ggatctgggg ctgttgatac tcaagcggcg aagagctcag cgaacattat 1467
Trp

ggcgaactgt ttctggctga aaaccctcct ctta 1501

<210> 2

<211> 439

<212> PRT

<213> Ochrobactrum anthropi IA NCIMB 41129

<400> 2

Met Gln Thr Pro Leu Ser Leu Arg Glu Arg Asp Ala Arg Val Ile Ala
1 5 10 15

Glu Ile Gly Arg Leu Arg Phe Ser Pro Leu Ser Leu Ile Gly Gly Lys
20 25 30

Gly Asn Arg Leu Ile Glu Glu Gly Gly Arg Ser Ile Leu Asp Leu Ser
35 40 45

Gly Ser Ala Gly Pro Ala Ala Leu Gly Tyr Gly His Pro Ala Ile Val
50 55 60

Glu Ala Val Glu Lys Ser Val Arg Asp Met Ala Gly Ala Ser Leu Leu
65 70 75 80

Leu Tyr Pro Asn Glu Ala Ala Val Ser Leu Ala Glu Asp Leu Leu Arg
85 90 95

Ile Thr Pro Gly Asn Gly Glu Arg Arg Val Trp Phe Gly His Ser Gly
100 105 110

Ser Asp Ala Asn Asp Cys Ala Val Arg Val Leu Thr Ala Ala Thr Lys
115 120 125

Arg Ser Arg Ile Ile Ser Phe Ile Gly Ser Tyr His Gly Asn Leu Thr
130 135 140

Gly Ser Met Gly Ile Ser Gly His Thr Ala Met Thr His Thr Leu Pro
145 150 155 160

Arg Pro Gly Val Leu Leu Leu Pro Tyr Pro Asp Pro Phe Arg Pro Arg
165 170 175

Phe Ser Ala Glu Ala Val Leu Glu Leu Leu Asp Tyr His Phe Ala Thr
180 185 190

Ser Cys Pro Pro Glu Gln Val Ala Ala Val Phe Ile Glu Pro Ile Leu
195 200 205

Ser Asp Gly Gly Leu Val Val Pro Pro Pro Ala Phe Leu Glu Ala Leu
210 215 220

Gln Asp Arg Cys Arg Lys His Gly Ile Leu Val Val Val Asp Glu Val
225 230 235 240

Lys Val Gly Leu Gly Arg Thr Gly Leu Met His Cys Phe Gln His Glu
245 250 255

Gly Leu Glu Pro Asp Met Val Val Phe Gly Lys Gly Leu Gly Gly Gly
260 265 270

Leu Pro Leu Ser Ala Val Val Gly Pro Gln Trp Val Met Asp His Ala
275 280 285

Pro Ala Phe Val Leu Gln Thr Thr Ala Gly Asn Pro Val Ala Thr Ala
290 295 300

Ala Gly Arg Ala Val Leu Asn Thr Ile Glu Arg Gln Gly Leu Ala Gln
305 310 315 320

Arg Ser Glu Arg Val Gly Gly Ile Phe Ala Asp Arg Leu Arg Arg Leu
325 330 335

Ser Asp Lys His Ser Ile Ile Gly Asp Val Arg Gly Arg Gly Leu Ala
340 345 350

Ile Gly Val Asp Leu Val Ser Asp Arg Gly Ser Arg Glu Pro Ala Pro
355 360 365

Val Thr Thr Thr Ala Lys Ile Ile Tyr Arg Gly Tyr Gln Leu Gly Ala
370 375 380

Ala Phe Thr Tyr Val Gly Leu Asn Ala Asn Val Leu Glu Phe Met Pro
385 390 395 400

Pro Leu Thr Leu Thr Glu Pro Glu Ile Asp Glu Ala Ala Asp Ile Val
405 410 415

Asp Gln Ala Ile Gly Asp Val Leu Asp Gly Lys Val Ala Asp Ser Asp
420 425 430

Val Ala His Phe Met Met Trp
435

<210> 3

<211> 436

<212> DNA

<213> Arthrobacter nicotianae NCIMB 41126

<400> 3

cggcgtgctt aacacatgca agtcgaacga tgatccccag cttgctgggg ggattagtgg 60

cgaacgggtg agtaacacgt gagtaacctg cccccgactc tgggataagc ccgggaaact 120

gggtctaata ccggatattc acttccttcc gcatggaggg tggtggaaag gtcttttccg 180

gtcggggatg gactcgcggc ctatcagctt gttggtgggg tagtggccca ccaaggcgac 240

gacgggtagc cggcctgaga gggtgaccgg ccacactggg actgagacac ggcccagact 300

cctacgggag gcagcagtgg ggaatattgc acaatgggcg caagcctgat gcagcgacgc 360

cgcgtgaggg atgacggcct tcgggttgta aacctctttc agtagggaag aagcgaaagt 420

gacggtacct gcagaa 436

<210> 4

<211> 400

<212> DNA

<213> Agrobacterium rhizogenes Bi NCIMB 41128

<400> 4

ggcggcaggc ttaacacatg caagtcgagc gccccgcaag gggagcggca gacgggtgag 60

taacgcgtgg gaatctaccc ttttctacgg aataacgcag ggaaacttgt gctaataccg 120

tatacgccct tcgggggaaa gatttatcgg gaaaggatga gcccgcgttg gattagctag 180

ttggtggggt aaaggcctac caaggcgacg atccatagct ggtctgagag gatgatcagc 240

cacattggga ctgagacacg gcccaaactc ctacgggagg cagcagtggg gaatattgga 300

caatgggcgc aagcctgatc cagccatgcc gcgtgagtga tgaaggccct agggttgtaa 360

agctctttca ccggtgaaga taatgacggt aaccggagaa 400

<210> 5

<211> 381

<212> DNA

<213> Agrobacterium rhizogenes Na NCIMB 41127

<400> 5

caggcttaac acatgcaagt cgagcgcccc gcaaggggag cggcagacgg gtgagtaacg 60

cgtgggaatc tacccttttc tacggaataa cgcatggaaa cgtgtgctaa taccgtatga 120

gcccttcggg ggaaagattt atcgggaaag gatgagcccg cgttggatta gctagttggt 180

ggggtaaagg cctaccaagg cgacgatcca tagctggtct gagaggatga tcagccacat 240

tgggactgag acacggccca aactcctacg ggaggcagca gtggggaata ttggacaatg 300

ggcgcaagcc tgatccagcc atgccgcgtg agtgatgaag gccctagggt tgtaaagctc 360

tttcaccggt gaagataatg a 381

<210> 6

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 6

gcctcacata tgcaaacacc gctttcattg cg 32

<210> 7

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 7

gcctcacccg ggttaccaca tcataaaatg ggcgacatc 39

Att ata tat cgc ggc tat caa ctc ggc gca gct ttt acc tat gtc ggc 1267
Ile Ile Tyr Arg Gly Tyr Gln Leu Gly Ala Ala Phe Thr Tyr Val Gly
375 380 385 390

ctc aat gcc aat gtt ctg gaa ttc atg ccc ccg ttg act ttg acc gag 1315
Leu Asn Ala Asn Val Leu Glu Phe Met Pro Pro Leu Thr Leu Thr Glu
395 400 405

ccg gaa ata gac gaa gca gcg gac atc gtt gat cag gcc att ggc gat 1363
Pro Glu Ile Asp Glu Ala Ala Asp Ile Val Asp Gln Ala Ile Gly Asp
410 415 420

gtt ctg gat ggg aag gtg gct gat agc gat gtc gcc cat ttt atg atg 1411
Val Leu Asp Gly Lys Val Ala Asp Ser Asp Val Ala His Phe Met Met
425 430 435

tgg taa ggatctgggg ctgttgatac tcaagcggcg aagagctcag cgaacattat 1467
Trp

ggcgaactgt ttctggctga aaaccctcct ctta 1501

Annexe 6

Sequence listings of oligonucleotides

[SEQ ID: No.6]

gcctcacata tgcaaacacc gctttcattg cg 32

[SEQ ID: No.7]

gcctcacccg ggttaccaca tcataaaatg ggcgacatc 39